FEATURE ARTICLE

APPLICATION MISMATCHES IN THE LABOUR MARKET IN SINGAPORE: ESTIMATES AND EFFECTS

INTRODUCTION

This study examines the effect of mismatches between jobseekers and the jobs that they applied for at the point of application. Using data from MyCareersFuture.sg, we study the effect of application mismatches in the labour market along six dimensions, namely – (i) industry, (ii) occupation, (iii) education, (iv) experience, (v) salary, and (vi) skills mismatches. The key factors driving application mismatch are identified, and machine learning techniques are used to estimate the effects of mismatch on long term unemployment, as well as the probability of unemployed jobseekers securing employment.

FINDINGS

FINDING 1
Of the six dimensions of mismatch identified, skills mismatch is the primary driver of application outcomes.

FINDING 2
A lower skills mismatch translates to higher application quality, and is thus associated with better application outcomes.

FINDING 3
Higher application quantity and lower duration between applications are associated with better application outcomes.

FINDING 4
Higher employer activity is associated with better application outcomes.

POLICY TAKEAWAY

The effects of skills mismatch, as well as jobseekers' application behaviour and employers' activity on labour market outcomes, suggest that there is scope to reduce information asymmetry and hence "search frictions". In particular, providing transparent and direct information on skills as well as job roles on MyCareersFuture.sg may help to improve job matching between employers and jobseekers.

EXECUTIVE SUMMARY

With job search activities increasingly migrating online, researchers now have new avenues to study the labour market, especially the job matching process, at an unprecedented level of detail. Using data from MyCareersFuture.sg, we study the different dimensions of mismatches between jobseekers and the jobs they applied for at the point of application, and find that skills mismatch is the most significant mismatch dimension. Segmenting skills mismatch into generic and non-generic skills mismatch, we find that non-generic skills mismatch is a more significant factor than generic skills mismatch.

Next, we apply machine learning approaches to identify the factors that have the greatest effect on labour market application outcomes. The two outcomes examined are the likelihood of unemployed jobseekers securing a successful job match and the risk of unemployed jobseekers becoming long-term unemployed. Our findings again suggest that skills mismatch is an important driver of these outcomes. Apart from skills mismatch, other factors that have a significant impact on application outcomes include the job search effort on the part of the jobseeker.

The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the Ministry of Trade and Industry or the Government of Singapore.[1]

INTRODUCTION

There are many different forms of labour market mismatches. In this study, we focus narrowly on the mismatches between the jobseekers and the jobs that they applied for at the point of application. While this was difficult to quantify in the past due to a lack of data, we can now leverage on the growing popularity of online job platforms like MyCareersFuture.sg to study application mismatches in the labour market at a more granular level.

This study will offer insights into six dimensions of application mismatches, namely – (i) industry, (ii) occupation, (iii) education, (iv) experience, (v) salary, and (vi) skills mismatches. A key contribution of this analysis is to quantify these mismatches, and empirically determine which of them are the main drivers of application mismatches in the labour market in Singapore. We then extend the analysis by using machine learning techniques to identify the factors that affect the likelihood of a successful job match among unemployed jobseekers, as well as the risk of unemployed jobseekers becoming long-term unemployed (LTU).[2]

The rest of the article is organised as follows. We start by providing a brief introduction to the academic literature in this field. We then describe the data, before presenting our methodology and results. The last section concludes.

LITERATURE REVIEW

The importance of skills as a driver of labour market outcomes is well recognised in the literature. Earlier work in the skills-biased technical change literature (Katz & Murphy, 1992; Autor, Katz, & Krueger, 1998) examined the divergent effect of skills on wages arising from technological change. More recent work on skills finds evidence of a positive correlation between wages and the groups of skills demanded by firms, with the difference in skills requirements accounting for around 12 per cent of the variation in wages (Deming & Kahn, 2017). In a similar vein, this study aims to provide a more granular look at the role of skills by studying its impact on labour market outcomes during the job application phase.

1 *We would like to thank Yong Yik Wei and Kuhan Harichandra for their useful suggestions and comments. We would also like to thank Workforce Singapore (WSG) for providing inputs related to MyCareersFuture.sg. All remaining errors belong to the authors.*
2 *For this study, an individual is defined as a LTU if he/she is actively searching for work and has been unemployed for six or more consecutive months.*

With the increasing availability of micro data on the job matching process (e.g., through online job portals), studies overseas have applied machine learning techniques on these datasets to identify the factors that are good predictors of various labour market outcomes. For example, Matty et al. (2012) estimated the risk of an individual becoming LTU by applying predictive modelling on UK administrative data, and found that close to 60% of the variation in LTU statuses could be explained by the model. Likewise, Chalfin et al. (2016) applied machine learning techniques on job application and demographic data in order to identify the factors that would be good predictors of job fit for applicants applying to become police officers. In the same vein, for this study, we apply machine learning approaches on data from MyCareersFuture.sg to tease out the factors that affect various labour market application outcomes such as the likelihood of a successful job match.

DATA

Jobs Bank was launched in 2014 as part of the Singapore Government's effort to strengthen the Singaporean core in the workforce. As part of the Fair Consideration Framework -- which mandates that companies consider Singaporeans fairly for job opportunities through an open, merit-based and non-discriminatory recruitment process, companies that wish to hire foreign employees are legislatively required to advertise vacancies on the Jobs Bank before they can be granted Employment Passes. Since April 2018, the Jobs Bank has been replaced by MyCareersFuture.sg (MCF), which aims to provide jobseekers with a smarter job search service through a skills-based job search functionality.

Traditionally, macro-surveillance and research involving Singapore's labour market have relied on data from surveys and administrative sources. With job postings and job search increasingly moving online, the MCF offers a compelling supplementary data source that can be tapped on to provide insights on demand and supply flows in the labour market on a real-time basis.

This study uses anonymised data on jobseekers as well as data on job postings from MCF. The dataset is augmented with firm-level data from the Accounting & Corporate Regulatory Authority's (ACRA) business registry, as well as employment-related data from government agencies. The data covers the period of January to December 2016.[3]

METHODOLOGY AND RESULTS

We first quantify the various dimensions of labour market application mismatches, and then estimate the extent to which they contribute to overall application mismatch on the MCF. Next, we employ machine learning techniques to identify the factors (including the different mismatch dimensions) that have the greatest impact on (i) the likelihood of a successful job match among unemployed jobseekers, and (ii) the risk of unemployed jobseekers becoming LTU.

Quantify the Dimensions of Application Mismatch

We disaggregate the concept of labour market application mismatch into six dimensions – (i) industry, (ii) occupation, (iii) education, (iv) salary, (v) experience, and (vi) skills. Traditionally, the first three mismatch dimensions tend to be more widely studied using established taxonomies (i.e., SSIC, SSOC and SSEC[4] respectively). However, the richness of the data from MCF allows us to study three other dimensions of mismatch. First, as the MCF dataset provides information about each individual's last drawn salary and also a job's offered salary, the two can be put together to analyse the extent of salary expectations mismatch.[5] Second, as we are able to obtain an individual's prior work experience from his/her resume, the work experience can be compared against the requirements of the job that he/she has applied for in order to measure the extent of experience mismatch.[6]

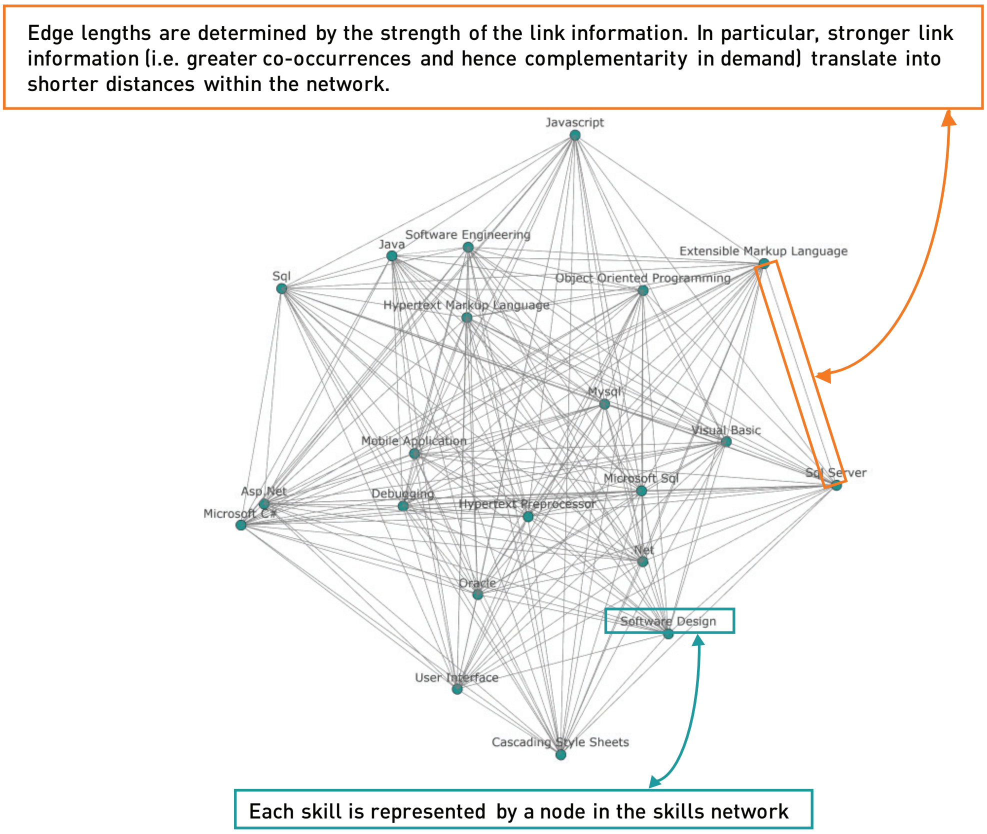
3 *While the MCF presents a unique opportunity for us to study the labour market at a granular level, it should be noted that the dataset is largely based on self-reported information from both jobseekers (e.g., last drawn salary and skills possessed) and employers. In addition, although the MCF accounts for a sizeable portion of the online job market in Singapore, there is likely to be some degree of self-selection among jobseekers and employers who opt to use the platform.*
4 *SSIC, SSOC and SSEC refer to the Singapore Standard Industrial Classification, Singapore Standard Occupational Classification and Singapore Standard Educational Classification respectively.*
5 *Jobs posted without an offered salary were dropped from our dataset.*
6 *Experience mismatch is measured by taking the difference between the years of experience demanded by a job, and the years an individual had spent in a relevant industry.*

Third, as the job postings and jobseeker resumes on MCF contain free-text descriptions of the skills that are required in jobs and the skills possessed by jobseekers respectively, we are able to apply text and network analytics to the MCF data to identify these skills and quantify the mismatch in skills between the jobseekers and the jobs they applied for. Specifically, this entails the following steps:

(i) Build a Skills List: We first train a natural language processing (NLP) algorithm to identify skills from free-text, and apply this to job descriptions in the job postings on MCF. This forms our base list of skills, which we then supplement with skills information provided by WSG and also from SkillsFuture Singapore's (SSG) Skills Frameworks for the various sectors.[7] By combining these sources of information, we are able to derive a consolidated list of approximately 1,000 skills.

(ii) Construct a Skills Network: Using the skills list compiled, we construct a network to map all pairwise relationships between the skills based on their co-occurrences in job descriptions (i.e., a skill is said to co-occur with another skill if they are both cited together in a job description). This approach is similarly employed by Djumalieva and Sleeman (2018), who applied their network model to detect communities of skills. In our network, the skills are represented by nodes, and every pairwise skills relationship is represented by an edge. The weight of each edge is determined by the extent of co-occurrences of each skill pair throughout our sample of job descriptions. Skill pairs that co-occur frequently are closer to each other in the network, and could be deemed to be more complementary in nature. A visualisation of this network is provided in Exhibit 1:

Exhibit 1: Visualisation of the Skills Network



7 *At the point when the research was conducted, SSG had published skills frameworks for the following 13 sectors: Accountancy, Aerospace, Early Childhood Care and Education, Electronics, Energy and Chemicals, Food Services, Hotel and Accommodation Services, Infocomm Technology, Logistics, Marine and Offshore, Precision Engineering, Retail, and Sea Transport.*

(iii) Segment Generic/Non-Generic Skills: Using the constructed skills network, we then compute the hub and entropy scores for each skill. Hub scores reflect the genericity of any given skill by measuring its propensity to be linked to all other skills in the network, while entropy scores measure how well each skill is spread out across the industries. We define a skill to be generic if it is well-connected to all other skills (i.e., high hub score) and also well spread out across industries (i.e., high entropy score). In particular, we apply a density-based clustering algorithm on the hub and entropy scores of all the skills in our network to select a set of skills with the most generic characteristics. Through this, we are able to classify the skills in our consolidated skills list into generic and non-generic skills. Exhibit 2 provides an illustration of the generic and non-generic skills across a sample of industries.

Exhibit 2: Generic/Non-Generic Skills by Industry

	Generic	Non-Generic	
Information & Communications (I&C)	• Information Technology • Data Analysis	• Java • Python • C++ • Javascript	• Hadoop • Node.JS • Microsoft SQL
Finance & Insurance (F&I)	• Sales	• Tax • Insurance • Investment • Hyperion	• Business Development • Banking • Risk Management
Healthcare	• Communication • Interpersonal • Teamwork	• Nursing • Medicine	• Occupational Therapy • Physiotherapy
Accommodation & F&B	• Management • Communication • Microsoft Office	• Food Safety • Housekeeping	• Public Relations

(IV) Quantify Skills Mismatches: Finally, based on the skills taxonomy created, we extract the skills from the resumes of individual jobseekers and compare them with the skills required in the jobs that they have applied for. This allows us to assess the extent to which individual jobseekers have the skills (both generic and non-generic) that are required in the jobs that they have applied for.

Assessing Application Mismatches

Having quantified the various dimensions of application mismatch, we apply an ordered-probit model to estimate each dimension's contribution to overall application mismatch on the MCF platform. Specifically, we run an ordered-probit model using the jobseeker's application status on MCF as our dependent variable.

The application status can take on the values of "Shortlisted", "Received" or "Unsuccessful".[8] We then estimate the impact of the various mismatch dimensions on the status of the application, and compute their average marginal effects on each of the application statuses:

8 *There are three broad application statuses in the MCF: (i) Shortlisted (i.e., "Successful" or "Under Review"), (ii) Received, and (iii) Unsuccessful. On an ordinal scale, jobseekers are considered matched when their application is "Shortlisted", and the level of mismatch increases when moving from the application status of "Received" to "Unsuccessful".*

$$\text{Prob}(Y_i = y) = \Phi\,(\alpha \cdot X_i + \beta_1 \cdot \text{IndustryMismatch}_i + \beta_2 \cdot \text{OccupationMismatch}_i\ \beta_3 \cdot \text{EducationMismatch}_i + \beta_4 \cdot \text{SalaryExpectationMismatch}_i + \beta_5 \cdot \text{ExperienceMismatch}_{ii} + \beta_6 \cdot \text{GenericSkillMismatch}_i + \beta_7 \cdot \text{NongenericSkillMismatch}_i) \quad (1)$$

where X_i is a vector of control variables for applicant i,
and Y_i is an ordered categorical variable reflecting the application status for applicant i

The average marginal effects of the different mismatch dimensions on the probability of an unsuccessful application are shown in Exhibit 3. Of the mismatch dimensions, we observe that only the effects of occupation and skills mismatch are significant. We also observe that the magnitude of the estimated effects of the two types of skills mismatch (i.e., generic and non-generic) far outweigh any other dimension. Specifically, a one-unit increase in generic skills mismatch is, on average, associated with a 1.5 percentage-point (pp) increase in the probability of an unsuccessful application, while a similar increase in non-generic skills mismatch is associated with a 4.3pp increase in the probability of an unsuccessful application.

Exhibit 3: Average Marginal Effects of Ordered Probit Model of an Unsuccessful Application

Variable	Marginal Effect (pp)
Industry Mismatch	0.116
Occupation Mismatch	0.408***
Education Mismatch	0.092
Salary Expectation Mismatch	0.000
Experience Mismatch	0.002
Generic Skill Mismatch	1.468***
Non-generic Skill Mismatch	4.345***

*Note: *** p-value < 0.01*

Utilising Machine Learning to Predict Labour Market Application Outcomes

Two labour market application outcomes are examined in this section: (i) likelihood of successful job match among unemployed jobseekers[9], and (ii) risk of an unemployed jobseeker becoming LTU. To study the factors that have the greatest impact on these outcomes, we employ a range of machine learning techniques from logistic regression to random forest. Apart from the six dimensions of mismatches identified earlier, we also include other individual-, job- and firm-level characteristics, as well as additional features such as application activity (e.g., the number of applications made by each jobseeker) and employer activity (i.e., the level of activity on MCF of each employer), in our models.

Based on the optimal model selected,[10] we identify the factors that are best able to predict each of the two labour market application outcomes, viz. successful job match and risk of LTU among unemployed jobseekers. Using Partial Dependence Plots (PDPs) to measure the average marginal effect of each of the factors on the outcomes, we identify three major factors:

(i) Application Quality
Consistent with the finding in the previous section, we find that skills mismatch is a key driver of labour market application outcomes. From the PDPs in Exhibit 4, we can see that on average, a higher skills mismatch score is associated with a lower probability of a successful job match among unemployed jobseekers (Exhibit 4a) and a higher risk of them becoming LTU (Exhibit 4b). This suggests that the quality of the application (i.e., whether jobseekers are applying for jobs that they have the skills for) is an important factor that affects the labour market application outcomes of jobseekers.

9 Successful job matches refer to applications where the individual subsequently worked for the company he/she applied to. Successful job matches are identified through a combination of MCF data and administrative data.
10 Due to the heavily imbalanced dataset, we evaluate the models by examining their Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curve respectively, and select the model with the highest AUC.
11 PDP shows the average marginal effect of each factor on the dependent variable within the dataset used. In other words, it plots how changes in the value of the factor (x-axis) changes the mean prediction of the dependent variable (y-axis), holding the rest constant.

Exhibit 4a: Effect of Skills Mismatch on the Probability of a Successful Application

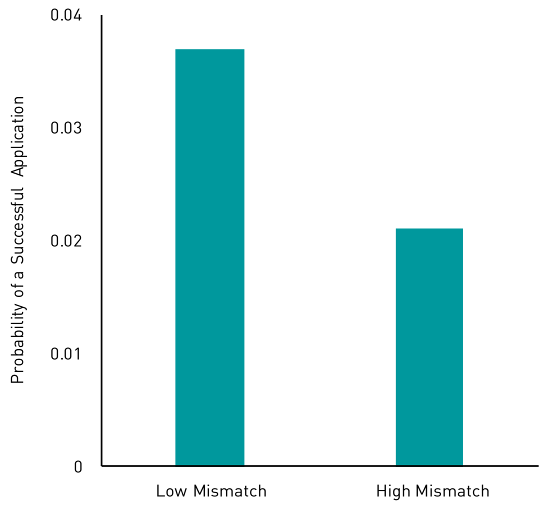
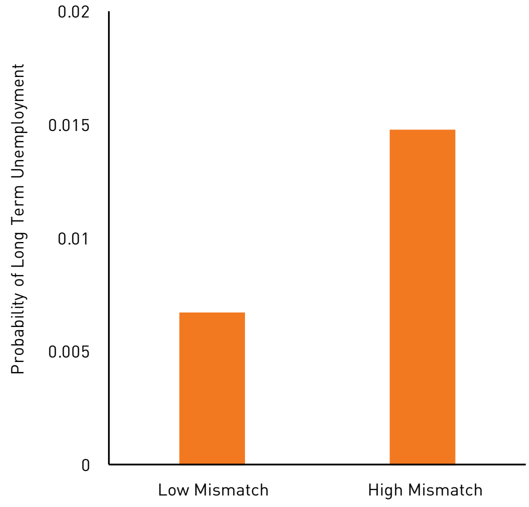


Exhibit 4b: Effect of Skills Mismatch on the Probability of Long Term Unemployment



(ii) Application Quantity and Time Duration

We also find that number of applications made by unemployed jobseekers and the duration between applications matter for labour market application outcomes (Exhibit 5). More specifically, we find that the higher the number of applications made by an unemployed jobseeker – which is a proxy for job search effort, the lower the likelihood that the jobseeker would become LTU (Exhibit 5b). This is especially so if the application made is of sufficient quality, as seen from the previous finding. Furthermore, we find that the shorter the time duration between the latest application and the previous application, the higher the probability that the unemployed jobseeker would have a successful job match (Exhibit 5a).

Exhibit 5a: Effect of Application Duration on the Probability of a Successful Application

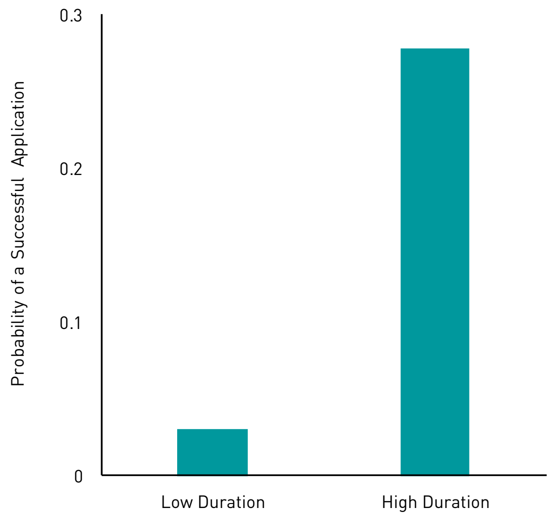
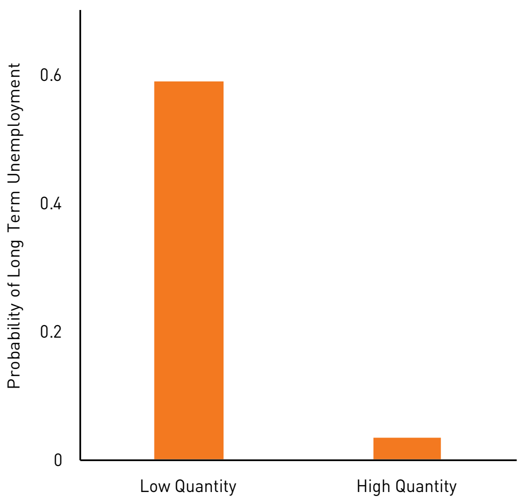


Exhibit 5b: Effect of Application Quantity on the Probability of Long Term Unemployment



(iii) Employer Activity

Lastly, we find that applying to jobs posted by employers that are active on the MCF platform also significantly improves labour market application outcomes (Exhibit 6).[12] For example, our results show that applying to employers who are active in reviewing applications on MCF raises the likelihood of a successful job match for the unemployed jobseeker and also reduces his/her risk of becoming LTU.

12 An employer's activity level is defined by the proportion of job applications to the employer that is under review by the employer.

Exhibit 6a: Effect of Employer Activity on the Probability of a Successful Application

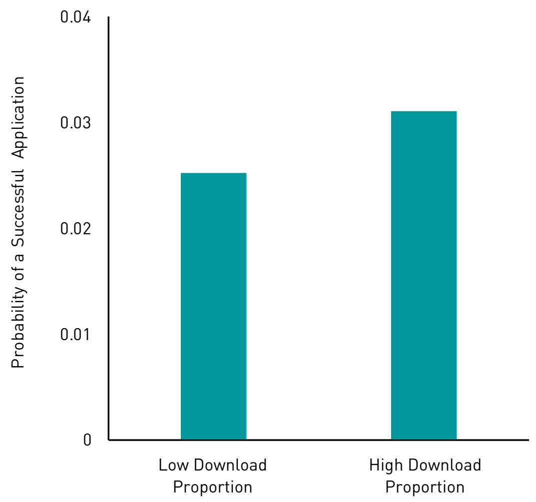
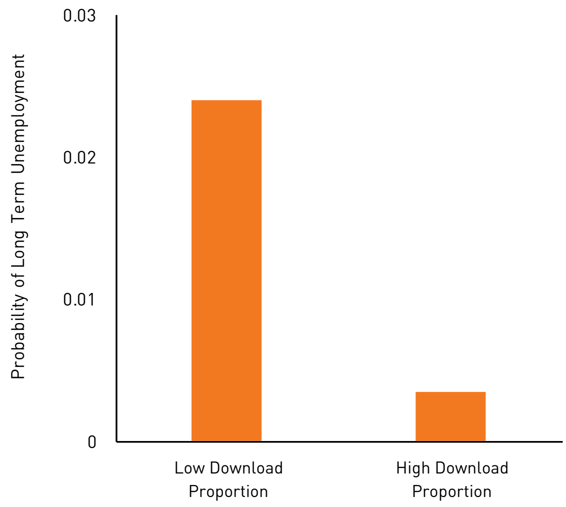


Exhibit 6b: Effect of Employer Activity on the Probability of Long Term Unemployment



SUMMARY AND CONCLUDING REMARKS

This study finds that skills mismatch is a key driver of labour market application outcomes. For instance, when jobseekers apply for jobs that they lack the skills for, which translates to applications of lower quality, they face a lower probability of a successful job match and a higher risk of becoming LTU. This finding suggests that it would be beneficial for jobseekers to be mindful of the skills requirements of the jobs that they are applying for, and to actively acquire these skills in order to improve their chances of success. To the extent that skills have been found to matter for application outcomes, this finding also suggests that WSG's efforts in encouraging employers to focus on skills to improve job fit may be seeing early signs of success.[13]

We also find evidence that the quantity of applications made by unemployed jobseekers and the time duration between applications for each jobseeker can be good predictors of application outcomes. In particular, unemployed jobseekers who make more applications appear to be at a lower risk of being LTU. Similarly, unemployed jobseekers with a shorter duration between their applications have a better chance of finding a successful job match. These findings suggest that there might be value in considering deliberate nudging of jobseeker behaviour on the MCF platform in order to improve application outcomes.

Lastly, we find that unemployed jobseekers who apply to employers that are more active on the MCF platform tend to enjoy better application outcomes. This finding has two broad implications. First, for jobseekers, it suggests that apart from understanding the requirements of the jobs that they are applying for, it is also pertinent for them to consider the hiring behaviour of employers. Second, for employers, it suggests that there is scope for those that are hitherto less active on MCF to become more active in order to improve their ability to find the right candidates via the platform.

More broadly, our study suggests that a greater availability of information on jobs (e.g., skills required for jobs) would be useful to reduce information asymmetry and hence "search frictions" for jobseekers, thereby improving application outcomes. In this regard, WSG's efforts to reduce information asymmetry on MCF, such as through more transparent and direct information on skills as well as job roles, can be seen as a positive step. We will continue to work with WSG to monitor the effectiveness of these initiatives in reducing search frictions, and find ways to continually improve job matching services on MCF.

Contributed by:

Mr Tek Yong Jian, Economist
Ms Jessica Foo, Economist
Economics Division
Ministry of Trade and Industry

13 *These efforts include encouraging employers to take part in programmes like the Adapt and Grow Professional Conversion Programmes and through the introduction of a job search functionality based on skills on MCF.*

REFERENCES

Autor, D., Katz, L., & Krueger, A. (1998). Computing Inequality: Have Computers Changed the Labor Market? *Quarterly Journal of Economics*, 1169-1241.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., & Mullainathan, S. (2016). Productivity and Selection of Human Capital with Machine Learning. *American Economic Review*, 124-127.

Deming, D., & Kahn, L. B. (2017). Skill Requirements across Frism and Labor Markets: Evidence from Job Postings for Professionals. *NBER Working Paper No. 23328*.

Djumalieva, J., & Sleeman, C. (2018). *Linking skills to occupations: Using big data to build a new occupational taxonomy for the UK. Nesta*.

Katz, L. F., & Murphy, K. M. (1992). Changes in Relative Wages, 1963-1987. *The Quarterly Journal of Econmoics*, 35-78.

Matty, S., Hart, A., Pinto, D., & Williamson, J. (2013). *Predicting likelihood of long term unemployment: the development of a UK jobseekers' classification instrument*. London: Department of Work and Pensions.